
**AFRL-OSR-VA-TR-2013-0534**

# (YIP-10) ROBUST NETWORK TRANSMISSION AND STORAGE USING CODING

**TRACEY HO**

**CALIFORNIA INSTITUTE OF TECHNOLOGY**

**08/09/2013**
**Final Report**



**AIR FORCE RESEARCH LABORATORY**
**AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSL**
**ARLINGTON, VIRGINIA 22203**
**AIR FORCE MATERIEL COMMAND**

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| 02-08-2013 | FINAL | 01 May 2010 - 30 Apr 2013 |

**4. TITLE AND SUBTITLE**

Robust Network Transmission and Storage Using Coding

**5a. CONTRACT NUMBER**

FA9550-10-1-0166

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Tracey Ho

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

California Institute of Technology
1200 El. California Blvd.
Pasadena, Calif 91125

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

USAF, AFRL
AF Office of Scientific Research
875 N. Randolph St. Room 3112
Arlington, VA 22203

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for Public Release: Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

This project focused on transmission and storage of information in networks, and considered robustness to adversarial errors, packet losses, link failure, mobility and topology dynamics. It established fundamental limits on performance (capacity, reliability and delay), as well as practical coding schemes and optimization techniques. Among our results, we developed network error correction theory for networks with non-uniform link capacities, non-multicast and a priori unknown number of errors. We applied this theory to design coding schemes for robust key distribution and streaming. We designed codes for error estimation, universal multicast codes robust to changes in network size and number of receivers, and error detection codes for distributed storage. We proposed efficient methods for obtaining network capacity bounds, characterized the impact of the failure of a single link on capacity of some families of networks, and showed equivalence between the Shannon capacity for saturated sources and the stable capacity of networks with probabilistic message arrivals. We characterized optimal resource allocation for maximizing the probability of data recovery under probabilistic access or failure of storage nodes, and optimizing transmission delay in disruption tolerant networks.

**15. SUBJECT TERMS**

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | Tracey Ho |
| U | U | U | UU | 34 | **19b. TELEPHONE NUMBER** *(include area code)* 626-395-4076 |

# Final report for AFOSR project "Robust Network Transmission and Storage Using Coding"

The work carried out under this grant established new theory as well as practical coding and optimization techniques for robust transmission and storage of information in networks. Our work studied fundamental limits on performance in terms of capacity, reliability and delay, and considered robustness to adversarial errors, packet losses, link failure, mobility and dynamically changing topology. The major findings are described below.

## 1 Coding for arbitrary/adversarial errors and network security

Coding for protection against errors in networks was introduced by Cai and Yeung [1, 2], and was extensively studied in the literature for the case of single-source multicast with a uniform error model, i.e. equal capacity network links/packets, any $z$ of which can be erroneous. The symmetry and simplicity of this case lead to simple cut set characterizations of capacity and coding techniques that do not extend straightforwardly to more general cases.

We developed a theoretical framework for error correction coding in more general network scenarios, showing new coding strategies and capacity bounding techniques that are quite different from the well-studied uniform single-source multicast case. This work extends network error correction to a much broader class of networks and provides novel achievability and converse techniques. The generality of the framework also opens up wider applicability of network error correction theory to new domains such as cryptography-based systems security and streaming codes, which are described further below.

Firstly, for the case of networks with nonuniform link capacities, we gave capacity bounds that account for the capacities of forward and feedback links on cuts, and connectivity between these links. This is in contrast to the uniform case where feedback links do not affect reliable information flow rate across a cut. We also devised novel coding schemes that tightly integrate error correction coding with partial error detection at intermediate nodes. These achievability and upper bounding results coincide in some cases. Our earlier work established results for the case of large capacity feedback links [3], and our work under this grant addressed the case of small capacity feedback links [4, 5] for which our previous bounds were loose.

Secondly, we considered the non-multicast case, for which determining capacity in general is an open problem even without errors. We showed how to combine cut set bounds for different sinks and error events to obtain tighter bounds on the error correction capacity region. We also showed a family of single- and two-source two-sink three-layer networks for which these bounds are tight, giving the exact capacity region [6]. An example of a three-layer network is shown in Figure 1. We extended this work to error/erasure correction coding for streaming data, described in the next section.

Thirdly, we designed rateless constructions of network error correction codes that can correct an a priori unknown number of errors by sending redundancy incrementally, unlike previous network error correction codes which are designed for a given number of errors. Our constructions are optimal in that receivers are able to decode when the amount of received and erroneous information

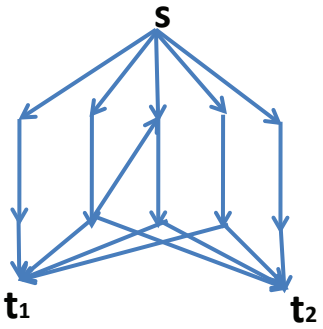


Figure 1: Example of a one-source two-sink network three-layer network for which we can find the exact capacity region.

satisfies the cut set bound with respect to the message size, with low complexity and overhead vanishing in the packet length. Specifically, our work under this grant [7] devised coding schemes that can exploit shared secret randomness or a low-rate secret channel between source and sink to reduce coding efficiency compared to our earlier construction in [8]. In the secret channel model, the source incrementally sends more linearly dependent redundancy of the source message through the network to combat erasures, and incrementally sends more linearly independent short hashes of the message on the secret channel to eliminate false information. The destination amasses both kinds of redundancy until decoding succeeds. In the shared randomness model, the source and destination share a small fixed random secret that is independent of the input message. Without a secret channel, both linearly dependent and independent redundancy have to be sent over the public and unreliable network, necessitating additional redundancy protection.

By removing the requirement for predetermined constraints on the number of errors, these rateless coding schemes can be combined with cryptographic signatures for security against adversarial errors. The combination of information theoretic coding with cryptographic operations is particularly useful in networks of computationally limited nodes such as low-power wireless nodes, which cannot perform complex cryptographic operations at a high rate. Our approach allows both redundant capacity and computation to be exploited as resources to achieve reliable communication rates higher than with either cryptographic or information theoretic approaches separately.

We also proposed new key agreement techniques for wireless networks in which one or more nodes may be adversarial and attempt to disrupt or compromise the key agreement process [9]. Our first scheme allows a pair of nodes to establish a common secret key using multiple multi-hop paths. Our secure error correcting code construction, designed for a specific topology, achieves better performance in terms of lower computational complexity and probability of error compared to our previous constructions in [10]. Our second scheme addresses the scenario of decentralized distribution of keys from a key pool. Each node needs a particular subset of the keys in the pool, and obtains them from the source and/or neighboring nodes who have already retrieved subsets of these keys. Our approach leverages our previously developed multisource network error correction codes [11] to achieve optimal resilience against errors introduced by adversarial nodes. Specifically, a node obtains coded combinations of its required keys from neighboring nodes that have subsets of these keys, achieving significantly stronger error resilience for a given redundancy overhead as compared to the case without coding.

## 2 Coding for unreliable links

### 2.1 Coding for streaming with packet erasures

In streaming data, information needs to be decoded by successive deadlines for uninterrupted playout at the receiver. We considered coding for packet erasures/errors in streaming of both stored and real-time (online) content. By modeling the streaming problem as a network error correction problem with a nested receiver structure, we were able to build on our work on non-multicast network error correction, described above, to analyze the streaming problem. We provided low-complexity achievable coding schemes, and, for various erasure/error models, converse bounds that match exactly or within a guaranteed ratio. These coding schemes do not rely on feedback, making them particularly suited for scenarios with broadcast and/or feedback delays. We considered different bursty and non-bursty erasure models, and showed significant differences in structural features of codes suited for these various models.

Specifically, for the case of stored content, i.e. all the content is initially present at the source, we studied the problem in which an arbitrary set of deadlines and demands can be specified. We first considered the problem of constructing codes that can correct any $z$ packet erasures (or errors), without a priori knowledge of which packets will be erased (erroneous). We showed that this problem could be modeled as a network error correction problem in which the receivers correspond to deadlines in the received packet stream by which particular pieces of information must be decoded, as illustrated in Figure 2. We characterized the capacity region of feasible demand vectors for any given set of deadlines and any $z$ erasures (errors), and provided a capacity-achieving coding scheme where no coding occurs across information demanded by different receivers [12]. We also considered a sliding window erasure model characterized by two parameters, erasure rate $p$ and a window size threshold $T$, in which the code is designed to correct erasure patterns where the number of erasures in any window of size at least $T$ is upper bounded by a fraction $p$ of the window size. We showed that our earlier coding scheme is approximately optimal for this erasure model also [13].

For the case of real-time streaming where messages are created at regular time intervals at a source, we studied the problem in which the receiver needs to decode each message within a given delay from its creation time, and considered three erasure models [14, 15]. In the first, a window-based erasure model, all erasure patterns containing a limited number of erasures in each sliding window of a specified length are admissible. In the second, a bursty erasure model, all erasure patterns containing erasure bursts of a specified maximum length separated by guard intervals of a specified minimum length are admissible. In the third, an i.i.d. erasure model, each transmitted packet is erased independently with a specified probability. We showed that a time-invariant intrasession code is asymptotically optimal over all codes (time-varying and time-invariant, intersession and intrasession) as the number of messages goes to infinity, for both the window-based erasure model and the bursty erasure model when the maximum erasure burst length is sufficiently short or long. For the bursty erasure model, we also showed that diagonally interleaved codes derived from specific systematic block codes are asymptotically optimal over all codes in certain other cases. For the i.i.d. erasure model, we derived an upper bound on the decoding probability for any time-invariant code, and showed that the gap between this bound and the performance of a family of time-invariant intrasession codes is small when the message size and packet erasure probability are small.

Besides streaming content, we also found another promising application of our online codes in decentralized control applications involving communication among a network of interacting stable (or individually feedback-stabilized) plants. Existing work on coding for decentralized control has primarily focused on stabilization of an unstable plant via communication over a noisy channel in

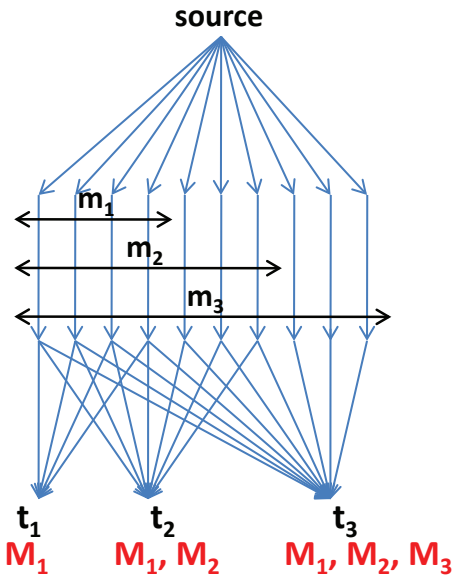


Figure 2: A single-source three-layer nested-network topology with three sinks, modeling a stored content streaming problem with deadlines $m_1, m_2, m_3$.

the feedback loop. In this setting, Sahai and Mitter [16] showed that older data must be recovered with increasing reliability (growing exponentially with delay), which is achieved by tree codes (Schulman [17]). In contrast, in some emerging domains such as smart grids, individual plants are stable while communication between plants is aimed at optimizing a cost or performance metric. Our online codes are designed for recovering timely rather than older data, and hence are suited for such stable distributed control problems [18].

### 2.2 Error-estimating codes

The concept of error-estimating codes (EEC) [19] was motivated by recent advances in wireless networking that leverage partially correct packets, for instance in scenarios such as rate-adaption [20, 21, 22] and real-time video streaming [23]. Unlike error correcting codes which correct errors, EEC allows the receiver to estimate the bit-error-rate (BER), using lower overhead compared to ECC. Using this BER information, the authors in [19] showed that the performance of upper-layer applications can be significantly improved. Furthermore, an error estimation code based on group sampling, which we term group-sampling error-estimating codes (GSEEC), was proposed in [19] and shown to achieve significantly lower communication overhead and computational complexity compared with existing error correcting codes.

We proposed in [24] a novel error estimation code, RAKEE, based on the theory of random walks. We provided theoretical analysis showing that RAKEE achieves the same asymptotic performance as GSEEC with respect to the packet length, that is constant communication overhead and linear coding complexity, while achieving better error decay performance. To be precise, under GSEEC the probability of unreliable estimation is proved to decay polynomially with increasing communication overhead, while under ALEEC such probability decays super-polynomially. Numerical experiments showed that RAKEE improves upon GSEEC in terms of both estimating bias and mean square estimation error.

## 3 Universal robust distributed multicast codes

Random linear network coding has been extensively studied for decentralized multicast [25]. Although robust to changes in topology and packet losses, existing schemes require knowledge of the network size and the number of sinks, or at least an upper bound. If these parameters are unavailable, such codes have no guarantees of correctness, hence they are not universal. Also, changing the field size to accommodate additional sinks or changes in network size entails changing the coding operation all nodes.

We developed the first universal distributed linear codes that have the advantage of not requiring a priori knowledge of network size and number of sinks, and being robust to changes in these parameters [26]. This is achieved by defining a hierarchical structure on the network that can be determined in a distributed fashion, and having each node choose coding coefficients randomly from a field of rational functions whose effective size grows with the distance from the source in this hierarchical structure. In particular, linear coding operations are chosen from finite subsets of an appropriate infinite field. A convenient field to use is the field of rational functions over $\mathbb{F}_2$. Operations over this field can be implemented via binary filters (convolutional codes) at each node. As information percolates down the network, each node makes its own estimate of the size of the subset of $\mathbb{F}_2(z)$ from which that node should choose its coding operations, so as to meet a pre-specified tolerance on the overall error-probability. We showed that this can be done using only information that can be percolated down the network at rates that are asymptotically negligible in the block-length, such that our codes are asymptotically rate-optimal. The code structure is designed to allow arbitrary changes in the topology and participating nodes, without requiring changes to existing random code choices. These codes also have polynomial-time design and implementation complexity.

## 4 Network capacity and impact of a single link

Characterizing the capacity region of a general non-multicast network is a major open problem in network coding. The complexity of existing computational methods for bounding the capacity of general networks grows exponentially with network size. This motivated us to investigate hierarchical methods for simplifying networks in order to find capacity bounds. We also studied the impact on network capacity of the loss of a single link in terms of the link capacity, as well as the effect of probabilistic arrivals of messages at source nodes. Our results are described in more detail below.

Firstly, we introduced in [27] a novel hierarchical approach for analyzing capacity regions of acyclic networks consisting of capacitated noiseless links with general demands. This approach sequentially replaces components of the network with simpler components containing fewer links or nodes, such that the resulting network is computationally simpler to analyze and its capacity provides an upper or lower bound on the capacity of the original network. The accuracy of the resulting bounds can be characterized as a function of the link capacities. Surprisingly, some families of network components can be simplified without affecting the network capacity.

Secondly, we studied the effect of loss of a single link of capacity $c$ on the capacity of a network of error-free bit pipes. We proved that if all the sources are available at a single source node, then removing a link of capacity $c$ cannot change the capacity region of the network by more than $c$ in each dimension [27]. We further extended this result to the case of multi-source, multi-sink networks for some special network topologies [28].

Thirdly, we considered the effect of probabilistic message arrivals and queuing on the capacity of general networks. The Shannon capacity is traditionally studied in the information theory/network coding literature. It is defined as the average rate of communicated information under the assump-

tion that the sources are saturated and can encode long blocks of source symbols, while receivers decode only after the entire block has been received. On the other hand, in many applications, the source messages arrive at source nodes statistically, resulting in idle and busy periods. The stable capacity of a network is defined as the set of all source arrival rate vectors that can be achieved by a stable solution in which each receiver node can eventually decode the desired source messages, and the queue size of each network node approaches a stable distribution over time.

Our work in [29] established an equivalence result between the Shannon capacity and the stable capacity of general non-multicast networks. Specifically, given a discrete-time network with memoryless, time-invariant, discrete-output channels, we proved that the Shannon capacity equals the stable capacity. This result applies even when neither the Shannon capacity nor the stable capacity is known for the given demands. The result also applies to both discrete alphabet channels and Gaussian channels.

## 5 Robust distributed storage

### 5.1 Distributed storage allocation

We investigated the problem of allocating a total storage budget $T$ across a number of distributed storage nodes so as to maximize recovery reliability. Specifically, the problem is to store a unit size data object using the given redundancy budget, such that the probability of recovering the data object is maximized under a given probabilistic access or failure model [30]. By using an appropriate code, successful recovery can be achieved whenever the total amount of data accessed is at least 1, the size of the original data object. This optimization problem is challenging in general because of its combinatorial nature, and a complete solution remains an open problem.

We studied several variations of the problem with different allocation models and access models. Among our results, we characterized a wide range of conditions of interest for which it is optimal to replicate a data object in entirety on a small number of nodes, or for which it is optimal to spread coded pieces of the data across many nodes.

Specifically, in the independent probabilistic access model, each storage node is accessed i.i.d. with a given probability $p$. We showed that the symmetric allocation that spreads the budget maximally over all nodes is asymptotically optimal in a regime of interest. Specifically, we derived an upper bound for the suboptimality of this allocation, and showed that when $p > 1/T$ the performance gap vanishes asymptotically as the total number of storage nodes grows. This is a regime of interest because it allows for a high probability of recovery. On the other hand, we showed that the symmetric allocation that spreads the budget minimally is optimal when $p$ is sufficiently small. In such an allocation, the data object is stored in its entirety in each nonempty node, making coding unnecessary. We also explicitly determined the optimal symmetric allocation (a practical family of allocations where all nonzero allocated values are equal) for a wide range of parameter values of $p$ and $T$, illustrated in Figure 3. Additionally, we derived a converse bound on the success probability, which is close to or coincides with the achievable performance for some parameter values, as shown in Figure 4.

In the fixed size subset access model, the objective is to maximize the probability of recovering the data object from a random subset of fixed size $r$. This problem is asymptotically equivalent to the fractional version, studied by Alon et al. [31], of a classical conjecture by Erdös on hypergraph matchings. We characterized a region of high recovery probability, in which the optimal allocation can be shown to allocate an amount $1/r$ to each of $\lfloor Tr \rfloor$ nodes.

We further built on this work to optimize message transmission delay using multiple paths in disruption tolerant networks. For minimization of expected delay we provided a complete characterization of the optimal symmetric allocation with respect to network parameter values [32]. We

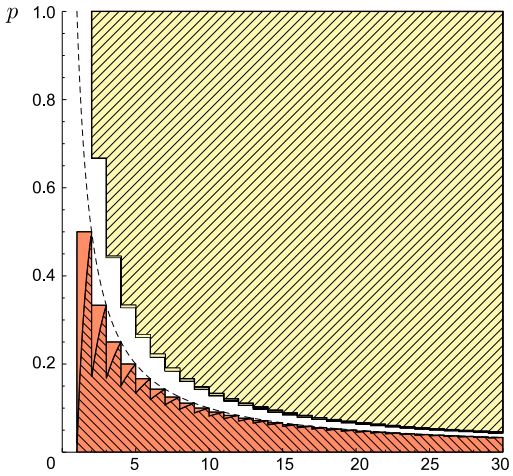


Figure 3: Plot of access probability $p$ against budget $T$. The black dashed curve marks the points satisfying $p = \frac{1}{T}$. Maximal spreading is optimal among symmetric allocations in the colored regions above the curve, while minimal spreading (uncoded replication) is optimal among symmetric allocations in the colored regions below the curve. In the remaining region near the curve, the optimal symmetric allocation changes in a complicated way due to integer effects.

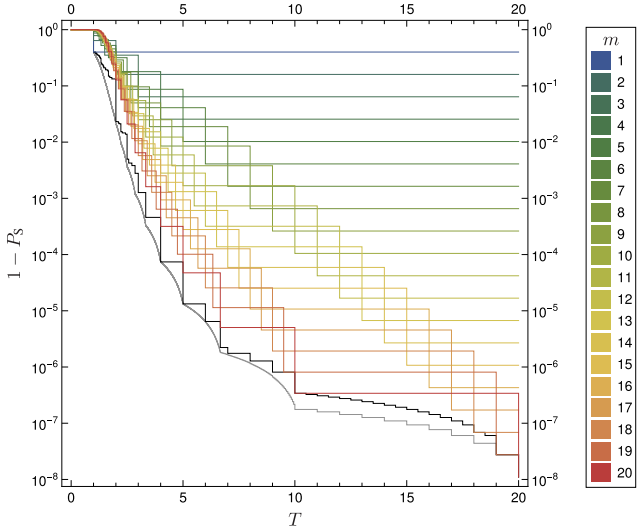


Figure 4: Plot of recovery failure probability against budget $T$ for each symmetric allocation for $(n, p) = \left(20, \frac{3}{5}\right)$. Parameter $m$ denotes the number of nonempty nodes in the symmetric allocation. The gray and black curves show two lower bounds for the recovery failure probability of an optimal allocation.

applied our results to design a data dissemination and storage protocol for mobile delay-tolerant networks, and showed in simulation experiments that the choice of storage allocation can have a significant impact on the recovery delay performance.

### 5.2 Detection of adversarial errors in distributed storage

We investigated in [33] the problem of maintaining an encoded dynamic coded distributed storage system where arbitrary adversarial errors can be introduced on an unknown subset of storage nodes. This distributed storage model had been introduced in [34] for the case without errors.

Leveraging the existing redundancy of the system, we proposed a simple linear hashing scheme to detect errors in the storage nodes. In particular, we showed that for a data object of total size $m$ using an $(n, k)$ MDS code, up to $t_1 = \lfloor (n-k)/2 \rfloor$ errors can be detected, with probability of failure smaller than $1/m$, by communicating only $O(n(n-k)\log m)$ bits to a trusted verifier. Our result constructs small projections of the data that preserve the errors with high probability and builds on a pseudorandom generator that fools linear functions.

## References


[1] R. W. Yeung and N. Cai, "Network error correction, part I: Basic concepts and upper bounds," *Commun. Inf. Syst.*, vol. 6, no. 1, pp. 19–36, 2006.

[2] N. Cai and R. W. Yeung, "Network error correction, part II: Lower bounds," *Commun. Inf. Syst.*, vol. 6, no. 1, pp. 37–54, 2006.

[3] S. Kim, T. Ho, M. Effros, and S. Avestimehr, "Network error correction with unequal link capacities," *Special issue of the IEEE Transactions on Information Theory dedicated to the scientific legacy of Ralf Koetter*, vol. 57, no. 2, pp. 1144–1164, 2011.

[4] T. Ho, S. Kim, Y. Yang, M. Effros, and A. S. Avestimehr, "On network error correction with limited feedback capacity," in *Information Theory and Applications Workshop (ITA), invited paper*, 2011.

[5] Y. Yang, T. Ho, S. Kim, M. Effros, and A. S. Avestimehr, "On network error correction with limited feedback capacity," 2012, preprint. [Online]. Available: http://www.its.caltech.edu/∼tho/yang.pdf

[6] S. Vyetrenko, T. Ho, and T. Dikaliotis, "Outer bounds on the error correction capacity region for non-multicast networks," in *Allerton conference on Communication, Control, and Computing*, September 2010.

[7] W. Huang, T. Ho, H. Yao, and S. Jaggi, "Rateless resilient network coding against byzantine adversaries," in *IEEE INFOCOM, submitted*, 2012.

[8] S. Vyetrenko, A. Khosla, and T. Ho, "On combining information-theoretic and cryptographic approaches to network coding security against the pollution attack," in *IEEE Asilomar Conference on Signals, Systems and Computers*, 2009.

[9] H. Yao, T. Ho, and C. Nita-Rotaru, "Key agreement for wireless networks in the presence of active adversaries," in *IEEE Asilomar Conference on Signals, Systems and Computers*, 2011.

[10] S. Jaggi, M. Langberg, T. Ho, and M. Effros, "Correction of adversarial errors in networks," in *Proc. of ISIT*, 2005.

[11] T. Dikaliotis, T. Ho, S. Jaggi, S. Vyetrenko, H. Yao, M. Effros, J. Kliewer, and E. Erez, "Multiple-access network information-flow and correction codes," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 1067–1079, February 2011.

[12] O. Tekin, S. Vyetrenko, T. Ho, and H. Yao, "Erasure correction for nested receivers," in *Allerton conference on Communication, Control, and Computing*, September 2011.

[13] O. Tekin, T. Ho, H. Yao, and S. Jaggi, "On erasure correction coding for streaming," in *Information Theory and Applications Workshop*, February 2012.

[14] D. Leong and T. Ho, "Erasure coding for real-time streaming," in *IEEE International Symposium on Information Theory*, 2012.

[15] D. Leong, A. Quereshi, and T. Ho, "On coding for real-time streaming under packet erasures," 2012, preprint. [Online]. Available: http://www.its.caltech.edu/∼tho/leongcoding.pdf

[16] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link, part i: Scalar systems," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3369 – 3395, Aug. 2006.

[17] L. J. Schulman, "Coding for interactive communication," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1745–1756, Nov. 1996.

[18] Y. Dong, T. Ho, N. Matni, and J. Doyle, "On coding for control of distributed systems," 2012, in preparation.

[19] B. Chen, Z. Zhou, Y. Zhao, and H. Yu, "Efficient error estimating coding: Feasibility and applications," Conference version was presented in SIGCOMM, technical report is avaible at: http://www.comp.nus.edu.sg/ ∼yuhf/eec-tr.pdf, Tech. Rep., 2010.

[20] J. C. Bicket, "Bit-rate selection in wireless networks," Master's thesis, MIT, 2005.

[21] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive mac protocol for multi-hop wireless networks," in *In Proc. of MOBICOM*, 2001.

[22] S. Wong, H. Yang, S. Lu, and V. Bharghavan, "Robust rate adaptation for 802.11 wireless networks," in *In Proc. of MOBICOM*, 2006.

[23] M. Elaoud and P. Ramanathan, "Adaptive use of error-correcting codes for real-time communication in wireless networks," in *In Proc. of INFOCOM*, 1998.

[24] H. Yao and T. Ho, "Error estimating codes with constant overhead: A random walk approach," in *IEEE International Conference on Communications*, 2011.

[25] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, October 2006.

[26] T. Ho, S. Jaggi, S. Vytrenko, and L. Xia, "Universal and robust distributed network codes," in *IEEE Infocom*, 2011.

[27] T. Ho, M. Effros, and S. Jalali, "On equivalence between network topologies," in *Allerton Conference on Communication, Control and Computing*, 2010.

[28] S. Jalali, M. Effros, and T. Ho, "On the impact of a single edge on the network coding capacity," in *Information Theory and Applications Workshop*, 2011.

[29] H. Yao, T. Ho, and M. Effros, "On the equivalence of shannon capacity and stable capacity in networks with memoryless channels," in *Proc. of ISIT*, 2011.

[30] D. Leong, A. Dimakis, and T. Ho, "Distributed storage allocations," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4733–4752, July 2012.

[31] N. Alon, P. Frankl, H. Huang, V. Rödl, A. Rucinski, and B. Sudakov, "Large matchings in uniform hypergraphs and the conjectures of Erdös and Samuels," *Journal of Combinatorial Theory Series A*, vol. 119, pp. 1200–1215, August 2012.

[32] D. Leong, A. Dimakis, and T. Ho, "Distributed storage allocations for optimal delay," in *IEEE International Symposium on Information Theory*, 2011.

[33] T. Dikaliotis, A. G. Dimakis, and T. Ho, "Security in distributed storage systems by communicating a logarithmic number of bits," in *IEEE International Symposium on Information Theory*, 2010.

[34] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. O. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," in *IEEE INFOCOM*, 2007.